

2.2.

Input: $\vec{a}^{(0)}$

*: here not considering activate function such as relu, sigmoid

$$(a) \vec{a}^{(1)} = W^{(1)} \vec{a}^{(0)} + \vec{b}^{(1)}$$

$$\vec{a}^{(2)} = W^{(2)} \vec{a}^{(1)} + \vec{b}^{(2)}$$

$$\text{Output: } \vec{a}^{(3)} = W^{(3)} \vec{a}^{(2)} + \vec{b}^{(3)}$$

(b) Input: $\vec{a}^{(0)}$

We know that Output = (a)'s output = $\vec{a}^{(3)} = W^{(3)} \vec{a}^{(2)} + \vec{b}^{(3)}$

\therefore to let (a) and (b) have same output

$$\vec{a}^{(3)} = W^{(3)} \vec{a}^{(2)} + \vec{b}^{(3)} = W^{(3)} (W^{(2)} \vec{a}^{(1)} + \vec{b}^{(2)}) + \vec{b}^{(3)} = W^{(3)} (W^{(2)} (W^{(1)} \vec{a}^{(0)} + \vec{b}^{(1)})) + \vec{b}^{(2)} + \vec{b}^{(3)}$$

\therefore based on the details above, let

$$\tilde{W} \vec{a}^{(0)} + \tilde{b} = W^{(3)} (W^{(2)} (W^{(1)} \vec{a}^{(0)} + \vec{b}^{(1)} + \vec{b}^{(2)} + \vec{b}^{(3)})$$

$$= W^{(3)} (W^{(2)} W^{(1)} \vec{a}^{(0)} + W^{(2)} \vec{b}^{(1)} + \vec{b}^{(2)} + \vec{b}^{(3)})$$

$$= W^{(3)} W^{(2)} W^{(1)} \vec{a}^{(0)} + W^{(3)} W^{(2)} \vec{b}^{(1)} + W^{(3)} \vec{b}^{(2)} + \vec{b}^{(3)}$$

$$= (W^{(3)} W^{(2)} W^{(1)}) \vec{a}^{(0)} + (W^{(3)} W^{(2)} \vec{b}^{(1)} + W^{(3)} \vec{b}^{(2)} + \vec{b}^{(3)})$$

$$\therefore \tilde{W} = W^{(3)} W^{(2)} W^{(1)}$$

$$\tilde{b} = W^{(3)} W^{(2)} \vec{b}^{(1)} + W^{(3)} \vec{b}^{(2)} + \vec{b}^{(3)}$$